



## Getting Started with the PIAAC Data

### Module Objectives

- Introduce users to the data collected through PIAAC and the resulting micro-level data files
- Describe the resources available to learn more about the survey data documentation and data files

### PIAAC Micro-level Data Files

- Organization for Economic Cooperation and Development (OECD) [International Public-use File \(International PUF\)](#)
- National Center for Education Statistics (NCES) [U.S. National Public-use File \(PUF\)](#)
- National Center for Education Statistics (NCES) [U.S. National Restricted-use File \(RUF\)](#)
- All data files contain
  - Responses to background questionnaire and cognitive assessment
  - Plausible values and weighting variables

### PIAAC Micro-level Data Files (Continued)

- Variable naming conventions are as follows
  - BQ variables: BQ section + Question # (e.g., B\_Q01a)
    - Trend items with IALS and ALL end in "\_T"
  - Cognitive items
    - Computer-based items begin with "C", "D", or "E"
    - Paper-based items begin with "P" (e.g., P924037R)
    - PS-TRE items begin with "U" (e.g., U01a000S)
- Naming conventions for derived variables vary

### PIAAC Micro-level Data Files (Continued)

- Data in the OECD and NCES PUFs that were suppressed for U.S. respondents include
  - Detailed earnings data
  - Occupational data, including detailed occupation codes, job title, and work responsibilities
  - Sampling selection variables

### OECD Micro-level Data Files

- [OECD International PUFs](#) are publicly available in SPSS and SAS formats for all countries except Australia and Cyprus
- Data are stored in separate files by country
  - Separate country files can be merged using the IEA International Database (IDB) Analyzer
- Data file names are stored as "prgXXXp1", where XXX is each country's three-letter country code
  - For example, "prgUSAp1" for the United States

### **NCES Micro-level Data Files**

- NCES U.S. (national) PUF and RUF are available in 3 formats
  - ASCII
  - SPSS
  - SAS
- The U.S. PUF and RUF include U.S.-only variables, in addition to all the variables in the International PUF

### **NCES Public-use Files**

- U.S.-only variables focus on: education, country of origin, language, race/ethnicity, training courses, occupation, health, work and earnings
- Include data for the U.S. routing of questions for those variables that differ from the international routing
- Background questionnaire items are named as follows
  - International variables adapted for the U.S. context end in "US" (e.g., D\_Q03US)
  - U.S.-only variables end in "USX" (e.g., B\_Q01bUSX)

### NCES Restricted-use Files

- Besides variables in the U.S. PUF, the U.S. RUF also contains variables suppressed in the PUFs
- The RUF contains responses for variables at the same level of detail as they were asked in the survey
- The RUF can be accessed through a [restricted-use license agreement](#) with NCES

### OECD PIAAC Data Documentation

Documentation for the OECD micro-data files includes

- [International technical report](#)
- [International data codebook](#)
- Data analysis manuals for [SAS](#) and [Stata](#)
- [SAS macros](#) and [Stata macros](#)
- Data compendia for [background variables](#) and [cognitive variables](#)

## International Technical Report

- [OECD Technical Report](#) of the Survey of Adult Skills describes
  - Assessment and Instrument Design
  - Platform Development
  - Field Operations and Quality Control
  - Sampling and Weighting
  - Data Analysis and Data Products
- The [NCES Technical Report](#) gives U.S.-specific details on sampling, field operations, and platform development

## International Codebook

- Describes the layout of the variables in the data file
- Documents values associated with response options for each survey questions

Sequence in Data	Name	Label	Type	Length	Width	Decimals	Domain	Value scheme	Missing scheme	Link to
1	CNTRYID	Country ID (ISO 3166, numeric)	Integer	Nominal	4	0	Sampling / weighting	Participating country codes (ISO 3166)	n Default ("blank"/Null), only	<a href="#">Values</a>
2	CNTRYID_E	Participating country or sub-national entity code (numeric)	Integer	Nominal	4	0	Sampling / weighting	Participating country or sub-national	e Default ("blank"/Null), only	<a href="#">Values</a>
3	SEQID	Sequential ID (randomly derived)	Integer	Nominal	5	0	Not assigned		Default ("blank"/Null), only	n/a
4	AGE_A	Person resolved age from BQ and QC check (derived)	Integer	Ratio	2	0	Sampling / weighting (derived)		Derived ("not stated" only)	<a href="#">Values</a>
5	GENDER_R	Person resolved gender from BQ and QC check (derived)	Integer	Nominal	1	0	Sampling / weighting (derived)	Gender	Derived ("not stated" only)	<a href="#">Values</a>
6	DISP_OBQ	Final disposition code for person - combining CI and BQ/RA (derived)	Integer	Nominal	2	0	Sampling / weighting (derived)	Disposition codes (CI, BQ, CBA)	Derived ("not stated" only)	<a href="#">Values</a>
7	DISP_MAIN	Final disposition code for person for Main task instrument (derived)	Integer	Nominal	2	0	Sampling / weighting (derived)	Disposition codes (PP, PRC, Global)	Derived ("not stated" only)	<a href="#">Values</a>
8	DISP_MAINWRC	Final disposition code for person for Main task instrument, including reading	Integer	Nominal	2	0	Sampling / weighting (derived)	Disposition codes (PP, PRC, Global)	Derived ("not stated" only)	<a href="#">Values</a>
9	A_NOT_F	Gender (Trend-IALS/ALL)	Integer	Nominal	1	0	Background questionnaire (trend)	Gender	BQ (numeric)	<a href="#">Values</a>
10	B_Q01a	Education - Highest qualification - Level	Integer	Nominal	2	0	Background questionnaire	ISCED 1997 PIAAC (1-16, incl non-formal)	BQ (numeric)	<a href="#">Values</a>
11	B_Q01a_T	Highest level of schooling (Trend-IALS/ALL)	Integer	Nominal	1	0	Background questionnaire (trend)	Trend - Highest level of schooling (1-4)	Trend - BQ (numeric) collapsed	<a href="#">Values</a>
12	B_Q01a1	Education - Highest qualification - Level of foreign qualification	Integer	Nominal	2	0	Background questionnaire	ISCED 1997 PIAAC (1-16, incl non-formal)	BQ (numeric)	<a href="#">Values</a>
13	B_Q01a1_C	Education - Highest qualification - Level of foreign qualification (collapsed, 1)	Integer	Nominal	2	0	Background questionnaire (derived)	Derived - ISCED 1997 PIAAC (1-16, incl non-formal)	BQ (numeric)	<a href="#">Values</a>
14	B_Q01b	Education - Highest qualification - Area of study	Integer	Nominal	2	0	Background questionnaire	ISCED Broad Fields of Education and Tr	BQ (numeric)	<a href="#">Values</a>
15	B_Q01c1	Education - Highest qualification - Age of finish	Integer	Ratio	2	0	Background questionnaire		BQ (numeric)	<a href="#">Values</a>
16	B_Q01c1_C	Education - Highest qualification - Age of finish (categorised, 6 categories)	Integer	Nominal	2	0	Background questionnaire (derived)	Derived - Age categories (1-6)	BQ (numeric)	<a href="#">Values</a>
17	B_Q01c1_T	Age at completion of highest level of schooling (Trend-IALS/ALL)	Integer	Ratio	2	0	Background questionnaire (trend)		BQ (numeric)	<a href="#">Values</a>
18	B_Q01c2	Education - Highest qualification - Year of finish	Integer	Scale	4	0	Background questionnaire		BQ (numeric)	<a href="#">Values</a>
19	B_Q01d	Education - Highest qualification - Month of finish	Integer	Nominal	2	0	Background questionnaire		BQ (numeric)	<a href="#">Values</a>
20	B_Q01d	Education - Highest qualification - Months elapsed since finished (DERIVED B)	Integer	Scale	3	0	Background questionnaire	Months	BQ (numeric)	<a href="#">Values</a>
21	B_Q01d_C	Education - Time elapsed since finished highest qualification (categorised, 5)	Integer	Ordinal	1	0	Background questionnaire (derived)	Derived - Time elapsed	BQ (numeric)	<a href="#">Values</a>
22	B_Q02a	Education - Current qualification	Integer	Nominal	1	0	Background questionnaire	Yes (1) / No (2)	BQ (numeric)	<a href="#">Values</a>
23	B_Q02a_T1	Education or training in last 12 months (Trend-IALS/ALL)	Integer	Nominal	1	0	Background questionnaire (trend)	Yes (1) / No (2)	Trend - BQ (numeric) collapsed	<a href="#">Values</a>
24	B_Q02a_T2	Courses toward certificate, diploma, or degree in program of studies in last 12	Integer	Nominal	1	0	Background questionnaire (trend)	Yes (1) / No (2)	Trend - BQ (numeric) collapsed	<a href="#">Values</a>
25	B_Q02b	Education - Current qualification - Level	Integer	Nominal	2	0	Background questionnaire	ISCED 1997 PIAAC (1-16, excl non-formal)	BQ (numeric)	<a href="#">Values</a>
26	B_Q02b_C	Education - Current qualification (collapsed, 10 categories)	Integer	Nominal	2	0	Background questionnaire (derived)	Derived - ISCED 1997 PIAAC (excl non-formal)	BQ (numeric)	<a href="#">Values</a>
27	B_Q02c	Education - Current qualification - Area of study	Integer	Nominal	2	0	Background questionnaire	ISCED Broad Fields of Education and Tr	BQ (numeric)	<a href="#">Values</a>
28	B_Q02b	Education - Uncompleted qualification	Integer	Nominal	1	0	Background questionnaire	Yes (1) / No (2)	BQ (numeric)	<a href="#">Values</a>
29	B_Q02b	Education - Uncompleted qualification - Level	Integer	Nominal	2	0	Background questionnaire	ISCED 1997 PIAAC (1-16, excl non-formal)	BQ (numeric)	<a href="#">Values</a>

### OECD PIAAC [SAS](#) and [Stata](#) Data Analysis Manuals

- Provide instructions for using the [SAS macros](#) and [Stata macros](#)
- Contain detailed explanations of the options and parameters in the macros
- Macros can be used to
  - Calculate statistically valid estimates of percentages, means, proficiency levels, percentiles
  - Conduct regression analysis

### OECD PIAAC Data Compendia

- Comprised of sets of tables that provide previously computed, weighted summary statistics for cognitive items and categorical background variables by country
- Support PUF users in gaining knowledge of PUF contents
- Useful for
  - Elaborating upon your hypothesis
  - Checking results of your own computations
  - Searching for variables
- Comparisons of cognitive item statistics across countries for reporting purposes is not appropriate

## OECD PIAAC Data Compendia (Continued)

- Information for categorical [background variables](#)
  - Sample size (Weighted N)
  - Weighted percent of responses by category
  - Percent missing, skipped, or refused responses
- Information for [cognitive items](#)
  - Sample size (Weighted N)
  - Weighted percent of correct and incorrect responses
  - Percent of respondents missing or not reached

## Sample Data Compendia Table

	Weighted N	Not Admin	Weighted N	Yes		No		Valid skip		Don't know		Refused		Not stated or inferred		Missing	
	All	%	Valid	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)	%	(SE)
Austria	5647341	1.83	5544140	63.33	(0.64)	36.67	(0.64)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	1.83	(0.16)
Belgium	4138042	5.15	3924851	66.19	(0.10)	33.79	(0.11)	0.00	(0.00)	0.00	(0.00)	0.02	(0.02)	0.00	(0.00)	5.15	(0.24)
Canada	23381067	0.00	23381067	68.59	(0.46)	31.38	(0.46)	0.00	(0.00)	0.01	(0.01)	0.00	(0.00)	0.02	(0.00)	0.00	(0.00)
Cyprus <sup>1,2</sup>	592296	17.69	487511	59.71	(0.80)	40.29	(0.80)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	17.69	(0.38)
Czech Republic	7395111	0.00	7395111	61.77	(0.42)	37.60	(0.43)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.62	(0.19)	0.00	(0.00)
Denmark	3629087	0.28	3618959	68.92	(0.43)	30.99	(0.44)	0.00	(0.00)	0.00	(0.00)	0.01	(0.00)	0.08	(0.03)	0.28	(0.06)
England	35422409	1.38	34932431	64.95	(0.38)	35.05	(0.38)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	1.38	(0.16)
Estonia	896163	0.37	892806	62.54	(0.50)	37.44	(0.50)	0.00	(0.00)	0.00	(0.00)	0.01	(0.00)	0.00	(0.00)	0.37	(0.06)
Finland	3496909	0.00	3496909	62.31	(0.67)	37.69	(0.67)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
France	40049569	0.44	39874109	57.60	(0.24)	42.30	(0.24)	0.00	(0.00)	0.03	(0.02)	0.00	(0.00)	0.27	(0.05)	0.44	(0.06)
Germany	53657540	1.48	52865188	68.64	(0.66)	31.36	(0.66)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	1.48	(0.16)
Ireland	2994368	0.37	2983331	55.20	(0.78)	44.76	(0.78)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.04	(0.03)	0.37	(0.10)
Italy	39369830	0.04	39353142	52.16	(0.39)	47.21	(0.40)	0.00	(0.00)	0.00	(0.00)	0.02	(0.02)	0.61	(0.15)	0.04	(0.04)
Japan	81059238	1.24	80057415	69.40	(0.33)	30.58	(0.33)	0.00	(0.00)	0.00	(0.00)	0.01	(0.01)	0.00	(0.00)	1.24	(0.12)
Korea	34602008	0.04	34587875	64.68	(0.61)	35.04	(0.62)	0.00	(0.00)	0.06	(0.04)	0.00	(0.00)	0.21	(0.06)	0.04	(0.03)
Netherlands	11160541	2.14	10921674	71.17	(0.61)	28.78	(0.60)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.06	(0.04)	2.14	(0.16)
Norway	3282755	2.17	3211433	71.33	(0.58)	28.62	(0.58)	0.00	(0.00)	0.00	(0.00)	0.02	(0.02)	0.04	(0.02)	2.17	(0.15)
Poland	26741987	0.00	26741987	55.05	(0.58)	44.90	(0.59)	0.00	(0.00)	0.03	(0.03)	0.02	(0.00)	0.00	(0.00)	0.00	(0.00)
Russian Federation <sup>3</sup>	87415088	0.00	87415088	55.60	(1.71)	44.36	(1.72)	0.00	(0.00)	0.00	(0.00)	0.05	(0.04)	0.00	(0.00)	0.00	(0.00)
Slovak Republic	3870993	0.25	3861337	57.38	(0.68)	42.60	(0.68)	0.00	(0.00)	0.00	(0.00)	0.01	(0.01)	0.01	(0.01)	0.25	(0.06)
Spain	31091563	0.74	30862627	53.38	(0.65)	46.57	(0.66)	0.00	(0.00)	0.02	(0.02)	0.00	(0.00)	0.03	(0.02)	0.74	(0.12)
Sweden	5985923	0.00	5985923	68.08	(0.56)	31.92	(0.56)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)	0.00	(0.00)
United States	203144374	3.98	195058736	68.06	(0.81)	31.64	(0.80)	0.00	(0.00)	0.02	(0.00)	0.00	(0.00)	0.28	(0.12)	3.98	(0.57)



### NCES PIAAC Data Documentation

- [U.S. Technical Report](#)
- U.S. [PUF](#) and [RUF](#) Codebooks
- [User's Guide for the NCES International Data Explorer \(IDE\)](#)
  - Provides instructions for using the PIAAC IDE, as well as detailed information on variables

### PIAAC Data Tools

- Web-Based Tools
  - [OECD International Data Explorer \(IDE\)](#) allows users to easily create statistical tables and charts
    - Contains national data files for all countries except Cyprus
  - [NCES International Data Explorer](#) is similar to the OECD IDE but includes U.S.-only variables in addition to the international variables
    - Contains U.S. national data as well as data for all countries except Australia and the Russian Federation
  - [NCES PIAAC Web Portal](#) provides ready-made tables and charts for the most common PIAAC data analyses

### PIAAC Data Tools (Continued)

- Micro-Data Tools
  - [International Database \(IDB\) Analyzer](#) (used with SPSS)
  - [SAS and Stata macros](#) developed by the OECD
  - [AM Statistical Software](#) developed by AIR
  - [R-package](#) developed by Statistics Austria (experimental)

### [NCES International Data Explorer \(IDE\)](#)

- Web-based tool that produces customizable tables and charts using PIAAC data
- A point-and-click method for producing custom results
  - Does not require any advanced statistical knowledge or software
  - Produces statistically valid results
- Can be used to calculate estimates of percentages, means, proficiency levels, and percentiles

### IDB Analyzer

- A stand-alone software program using SPSS as the computational engine for analyzing data from PIAAC
- Allows for creation of analysis datasets by
  - Combining specific country data files
  - Selecting specific subsets of variables
- Accounts for complex sample design with plausible values methodology, including correct calculation of standard errors
- Can be used to
  - Calculate estimates of percentages, means, proficiency levels, and percentiles
  - Conduct regression analyses

### **Linking PIAAC with IALS and ALL**

- Micro-data files containing data from the International Adult Literacy Survey (IALS) and the Adult Literacy and Life Skills Survey (ALL) can be ordered from Statistics Canada using the following links
  - [International Adult Literacy Survey micro-data files](#)
  - [Adult Literacy and Life Skills Survey micro-data files](#)
- These files contain results for literacy (IALS and ALL) and numeracy (ALL) on the same scales as the results from the Survey of Adult Skills (PIAAC)
- The process for linking results from PIAAC and those from IALS and ALL is described in Chapter 17.4.2 of the [OECD Technical Report](#)
- Both the NCES and OECD PIAAC IDEs contain trend data from IALS and ALL as well

### Summary

- Introduced users to the data collected through PIAAC and the resulting micro-level data files
- Described the resources available to learn more about the survey data documentation and data files

### Resources

- [International Public-use File \(International PUF\)](#)
- [U.S. National Public Use File \(PUF\)](#)
- [U.S. National Restricted Use File \(RUF\)](#)
- [Restricted-use license agreement](#)
- [OECD International Technical Report](#)
- [International Data Codebook](#)
- [SAS Data Analysis Manual](#)
- [Stata Data Analysis Manual](#)
- [SAS macros](#)
- [Stata macros](#)
- [Background variables](#)
- [Cognitive variables](#)
- [NCES U.S. Technical Report](#)
- [NCES International Data Explorer](#)
- [OECD International Data Explorer](#)
- [NCES PIAAC Web Portal](#)
- [International Database \(IDB\) Analyzer](#)
- [SAS and Stata macros](#)
- [AM Statistical Software](#)
- [R-package](#)
- [International Adult Literacy Survey micro-data files](#)
- [Adult Literacy and Life Skills Survey micro-data files](#)